

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

```
Run on:      March 1, 2001, 15:43:15 ; Search time 210.42 Seconds
              (without alignments)
              108.227 Million cell updates/sec
```

```

Title:      US-09-331-631A-1
Perfect score: 3542
Sequence:   1 MAINTSNLCSLFLFLSLFL.....SPRSTKQKQPLVSLDPVGF 6666

```

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%
	Listing first 45 summaries

```
Database :
_A_GeneSeq_36 : *
1: /SIDSL/gcgdata/geneseq/geneseqp/AA1980.DAT : *
2: /SIDSL/gcgdata/geneseq/geneseqp/AA1981.DAT : *
3: /SIDSL/gcgdata/geneseq/geneseqp/AA1982.DAT : *
4: /SIDSL/gcgdata/geneseq/geneseqp/AA1983.DAT : *
5: /SIDSL/gcgdata/geneseq/geneseqp/AA1984.DAT : *
6: /SIDSL/gcgdata/geneseq/geneseqp/AA1985.DAT : *
7: /SIDSL/gcgdata/geneseq/geneseqp/AA1986.DAT : *
8: /SIDSL/gcgdata/geneseq/geneseqp/AA1987.DAT : *
9: /SIDSL/gcgdata/geneseq/geneseqp/AA1988.DAT : *
10: /SIDSL/gcgdata/geneseq/geneseqp/AA1989.DAT : *
11: /SIDSL/gcgdata/geneseq/geneseqp/AA1990.DAT : *
12: /SIDSL/gcgdata/geneseq/geneseqp/AA1991.DAT : *
13: /SIDSL/gcgdata/geneseq/geneseqp/AA1992.DAT : *
14: /SIDSL/gcgdata/geneseq/geneseqp/AA1993.DAT : *
15: /SIDSL/gcgdata/geneseq/geneseqp/AA1994.DAT : *
16: /SIDSL/gcgdata/geneseq/geneseqp/AA1995.DAT : *
17: /SIDSL/gcgdata/geneseq/geneseqp/AA1996.DAT : *
18: /SIDSL/gcgdata/geneseq/geneseqp/AA1997.DAT : *
19: /SIDSL/gcgdata/geneseq/geneseqp/AA1998.DAT : *
20: /SIDSL/gcgdata/geneseq/geneseqp/AA1999.DAT : *
21: /SIDSL/gcgdata/geneseq/geneseqp/AA2000.DAT : *
```

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3542	100.0	666	19	W62828	Macadamia integrifolia
2	3412	96.3	666	19	W62829	Macadamia integrifolia
3	3227	91.1	625	19	W62830	Macadamia integrifolia
4	1127	31.8	566	13	R20181	Sequence encoded by
5	1103.5	31.2	590	19	W62832	Gossypium hirsutum
6	1025	28.9	525	19	W62831	Theobroma cacao an
7	898	25.4	593	19	W62835	Zea mays antimicro
8	864.5	24.4	637	19	W62837	Hordium vulgare an
9	849.5	24.0	626	20	V15244	Peanut allergen, A
10	849.5	24.0	626	20	V25657	Peanut allergen 11
11	849	24.0	614	18	W22149	Peanut allergen An
12	849	24.0	614	19	W62834	Arachis hypogaea a

15	148.5	4.2	911	18	W31866
14	148.5	4.2	911	20	W35957
44	153.5	4.3	1041	20	W30613
42	155.5	4.4	1382	18	W31867
41	155.5	4.4	968	20	W55966
40	156.5	4.4	395	17	W03474
39	157.5	4.4	515	15	R47128
38	158	4.5	1326	20	W55933
37	158.5	4.5	316	13	R26941
36	161.5	4.6	1233	20	Y55954
35	164.5	4.6	1135	21	Y55954
34	166	4.7	493	21	R68784
33	167	4.7	512	20	R26944
32	168	4.7	510	20	Y45912
31	171	4.8	1299	21	Y58633
30	180	5.1	499	9	F82755
29	180	5.1	472	15	R47127
28	190	5.4	360	17	W03627
27	191	5.4	412	17	W03626
26	195.5	5.5	1239	20	Y55931
25	195.5	5.5	562	16	R70491
24	211.5	6.0	1162	21	Y58500
23	216.5	6.1	1898	20	Y30795
22	238	6.7	611	20	Y29039
21	513.5	14.5	371	20	Y40914
20	700.5	19.8	409	20	W90342
19	707.5	20.3	415	20	W90340
18	756	21.0	444	20	W90340
17	806.5	22.8	484	20	W90341
16	808	22.8	524	20	W90339
15	831.5	23.5	605	20	W40999
14	840.5	23.7	605	19	W62838
13	843.5	23.8	626	18	W22150
12	845	23.8	626	18	W22150
11	845	23.8	626	18	W22150
10	845	23.8	626	18	W22150
9	845	23.8	626	18	W22150
8	845	23.8	626	18	W22150
7	845	23.8	626	18	W22150
6	845	23.8	626	18	W22150
5	845	23.8	626	18	W22150
4	845	23.8	626	18	W22150
3	845	23.8	626	18	W22150
2	845	23.8	626	18	W22150
1	845	23.8	626	18	W22150

ALIGNMENTS

Result	ID	Accession	Description	Location/Qualifiers
XX	W62828	standard; Protein; 666 AA.		
XX	W62828;			
XX	27-OCT-1998	(first entry)		
XX	Macadamia integrifolia	antimicrobial protein.		
XX	antimicrobial protein;	infestation; control.		
XX	Macadamia integrifolia.			
XX	Key			
XX	Peptide	1..28		
XX	FT	/note="signal peptide"		
XX	FT	29..666		
XX	FT	/note="mature protein"		
XX	W09827805-A1.			
XX	02-JUL-1998.			
XX	22-DEC-1997;	97WO-AU00874.		
XX	20-DEC-1996;	96AU-0004275.		
XX	(RETR-)	COOP RES CENT TROPICAL PLANT PATHOLOGY.		
XX	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP			
XX	WPI: 1998-377279/32.			
XX	N-PSDB: V42310.			

PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
XX
XX
PS Claim 1; Page 34-36; 96pp; English.
XX

CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
XX
XX

SQ Sequence 666 AA:

Query Match 100.0%; Score 3542; DB 19; Length 666;
Best Local Similarity 100.0%; Pred. No. 2e-299;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAINTSNLCSLFLSLFLSTVTSLAESFPDQREYEBCRQCMOLETSGQMRVCSQCD 60
DB 1 maintsnlcsllflslflstvtslaesefdrqyeecrkqcmqletsqgmtrcvsgcd 60
QY 61 KREEDIDMSKYDNOEDPQTECCQCRRCROESGPRQOQYCCORCKEICEEEYENRQR 120
DB 61 ktreedidmskydngedpqtccqqrtrcrqesgprqyqqrckelceeeeynrqr 120
QY 121 DPQOQYEQCQKRCQRETEPRHMQTCQRCERRYEKERRKQKRYEQQREDEKYEERH 180
DB 121 dpqoyeqcqhkgqrretprhmqtcqrccerryekekryeqqredeekyeerm 180
QY 181 KEEDNKRDPQREYEDCRRRCQEQEROHQCLRCREORHGRGDMNPNRGSGSGRY 240
DB 181 keednkrdpqreyedcrrrcqeqqrccrrcqrqyqgdmnprrgsgrgy 240
QY 241 EEGEESQSNPYFEDRSLSSTRFTEGHSVLENFYGRSKLLRALKNRYLVLEANPNA 300
DB 241 eegeesqsnpyfderlsstrfteeghsvlenfygrskllralknrylvleapna 300
QY 301 FVLPTHLDADAILLVYGRGALKMIHHDNRESYNLECGDVIRIPAGTTFYLLNRDNNERL 360
DB 301 fvlpthldadaillyvgrgalkmihhdnresynlecgdviripagctfyllnrddnerl 360
QY 361 HIAKFLQITSTPGQYKEFPAGQONPEPLSTFSKEILEALNTQTEKLRGVGQOREG 420
DB 361 hiakfqlitstpgykeffpagqonpepylstfskeilealntqteklrgvlgqoreg 420
QY 421 IIRASQEQIRELTRDSESRHWHIRRGESSSRGPYNLFNKRPYSNKYQAYEVKPEDYR 480
DB 421 iirasgeqireltrdssesrhwihrrgesssrgpynlfnkrrplysnkyqayevkpedyr 480
QY 481 QLODMLSLFIANYTQGSMMGPFRNTKSTKVYVYVAGSEADVEMACNLSGRHGRGGGR 540
DB 481 qlodmlslvfiannytqgsmmgpfntkstklyvvvvasgeadvemacnlsgrhgrgggr 540
QY 541 HEEEDVHVEQYARLSKREAVLVLAGHPVVFSSGENULLFAGINNQNNTENLAR 600
DB 541 heeedvhveqyarrlskreavlvlaghpvfvssgenullfaginnqnntenlar 600
QY 601 ERNVLOQIEPQAMELAFAPRKREVESEFSNODSLTFPGPROHQOQSPSTKQOQPLVSI 660
DB 601 ernvloqiepqamela faaprkrveesefsnodsltfpgprhqoqspstskqoqplvsi 660
QY 661 LDFVGF 666
DB 661 ldfvgf 666

RESULT 2
W62829 W62829 standard; Protein; 666 AA.

XX W62829;
XX
XX
DT 27-OCT-1998 (first entry)
XX

DE Macadamia integrifolia antimicrobial protein.
XX
XX antimicrobial protein; infestation; control.
XX
XX
OS Macadamia integrifolia.

XX
XX
FH Key Location/Qualifiers
FT peptide 1..28
FT /note= "signal peptide"
FT Protein 29..666
FT /note= "mature protein"

XX W09827805-A1.

XX 02-JUL-1998.

XX 22-DEC-1997; 97W0-AU00874.

XX 20-DEC-1996; 96AU-0004275.

XX (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.

XX Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;

XX WPI: 1998-377279/32.

XX N-PSDB; V42311.

PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals

XX Claim 1; Page 39-41; 96pp; English.

CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.

SQ Sequence 666 AA:

Query Match 96.3%; Score 3412; DB 19; Length 666;
Best Local Similarity 96.1%; Pred. No. 4.1e-288;
Matches 640; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 1 MAINTSNLCSLFLSLFLSTVTSLAESFPDQREYEBCRQCMOLETSGQMRVCSQCD 60
DB 1 maintsnlcsllflslflstvtslaesefdrqyeecrkqcmqletsqgmtrcvsgcd 60
QY 61 KREEDIDMSKYDNOEDPQTECCQCRRCROESGPRQOQYCCORCKEICEEEYENRQR 120
DB 61 ktreedidmskydngedpqtccqqrtrcrqesgprqyqqrckelceeeeynrqr 120
QY 121 DPQOQYEQCQKRCQRETEPRHMQTCQRCERRYEKERRKQKRYEQQREDEKYEERH 180
DB 121 dpqoyeqcqhkgqrretprhmqtcqrccerryekekryeqqredeekyeerm 180
QY 181 KEEDNKRDPQREYEDCRRRCQEQEROHQCLRCREORHGRGDMNPNRGSGSGRY 240
DB 181 keednkrdpqreyedcrrrcqeqqrccrrcqrqyqgdmnprrgsgrgy 240
QY 241 EEGEESQSNPYFEDRSLSSTRFTEGHSVLENFYGRSKLLRALKNRYLVLEANPNA 300
DB 241 eegeesqsnpyfderlsstrfteeghsvlenfygrskllralknrylvleapna 300
QY 301 FVLPTHLDADAILLVYGRGALKMIHHDNRESYNLECGDVIRIPAGTTFYLLNRDNNERL 360
DB 301 fvlpthldadaillyvgrgalkmihhdnresynlecgdviripagctfyllnrddnerl 360
QY 361 HIAKFLQITSTPGQYKEFPAGQONPEPLSTFSKEILEALNTQTEKLRGVGQOREG 420
DB 361 hiakfqlitstpgykeffpagqonpepylstfskeilealntqteklrgvlgqoreg 420
QY 421 IIRASQEQIRELTRDSESRHWHIRRGESSSRGPYNLFNKRPYSNKYQAYEVKPEDYR 480
DB 421 iirasgeqireltrdssesrhwihrrgesssrgpynlfnkrrplysnkyqayevkpedyr 480

Db	421	lisaqeqirfreltrddsserrwhlirigessrrpynlfnkrlpysnkyygagayewkpedyr	480
Qy	481	QlQMDLISVFIAANTQGSMMGPFFNTNSTKVVVVASGEADVENACPHLSGRHGGRGGR	540
Db	481	qlqmdvsvfianltqgsmmgpfntftrstkvvvasgeadvenacphlsgrhgyrgr	540
Qy	541	HEEEDVHYEOWRRLSKREAIIVYLAPHVVEVSSGNNELLTFPGINAOHNHNFAGR	600
Db	541	heeedvhyevkarlskreaiivpyvhpvfvssgennelltfghnaohnenflagr	600
Qy	601	ERNVLOQIEPQAMELAFAPRKEVEESFNQSDSIFPPGRHOQDSPRSTKQOQPLVSI	660
Db	601	ernvlgqepqamelaafaprkeveelfnsgdesiifppgrqhqqgsrstktqqlvsi	660
Qy	661	LDVGF 666	
Db	661	ldvgt 666	

RESULT	3
ID	W62830
ID	W62830 standard; Protein; 625 AA.
XX	
AC	W62830;
XX	
DT	27-OCT-1998 (first entry)
XX	
DE	Macadamia integrifolia antimicrobial protein.
XX	
KW	antimicrobial protein; infestation; control.
XX	
OS	Macadamia integrifolia.
XX	
FT	Key
FT	Location/Qualifiers
FT	Peptide
FT	1..28
FT	/note="signal peptide"
FT	29..666
FT	Protein
FT	/note="mature protein"
XX	
PN	W09827805-A1.
XX	
PD	02-JUL-1998.
XX	
PF	22-DEC-1997; 97MO-AU00874.
XX	
PR	20-DEC-1996; 96AU-0004275.
XX	
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX	
PI	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX	
DR	WPI; 1998-377279/32.
XX	
DR	N-PSDB; V42316.
XX	
PT	Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT	useful for controlling microbial infestations of plants or mammals
XX	
PS	Claim 1; Page 43-45; 96pp; English.
XX	
CC	The sequence is that of an antimicrobial protein which can
CC	be used to control microbial infestations in plants and mammalian
CC	animals.
XX	
XX	Sequence 625 AA;
XX	

Query Match	91.1%	Score 3227	DB 19	Length 625
Best Local Similarity	96.6%	Pred. No. 4.5e-272		
Matches 604	Conservative 9	Mismatches 12	Indels 0	Gaps 3
 42 OCMOLETSGOMRCVSCDQKFEEDIDMSRYDNOEDPTECCQCRRCROESGPRQOY 101				
db	1 qemleesgqmrtrvscgdkrlfeedidmsrydyndepqltecgcqqrtrcrgesdppqgy 60			

QY	102	CORCKEICEEEEXENNRDDQOQYEDQCKHRCORREPRHMQTCOORCERRERKEKRRKQ	161
Db	61	qtrrckelceeeeeeyntrdpdqyqeqdkrcqrtrcebrtmhqicqgrceeryekerktkq	120
QY	162	OKRYEEOORDEDEEKEERBMKEEDMKRPDQOQREYEDCRRRCEDQEPROHQHCOLRCRROOR	221
Db	121	qkryeeqqrredeeykyeermtkgadkkrdpqqrqeydcrrrheegeqeprrlyqyqqrtrcqeqqr	180
QY	222	OHGRGDMMNNPORGSGSYEBEGEEQSDNPPYFDEKSLSTRFTEGHSIVLENFYGRSK	281
Db	181	qhgqrgdlimpqrgsgsgyeegekgsdnpyfiderslstrfrceeghlsvlenfygrsk	240
QY	282	LLRAKKNRVLVLNNPAPVLPRLHDADALLVYGGRGALKMTHHNNRSTYLNCEGDVY	341
Db	241	llrralknrrvvlleannpavlprrlhdadallvlyggrrgalkmthdnrstynlcegdvyl	300
QY	342	RIPAGTFTVYLNRDNNERLHAKFLQRTSTPFGQYKEFFPFGGONPEPYLSTFKEITLFAA	401
Db	301	rtpagttftvylnrddnnelrhakflqrltstprgqykelfppgggnpemylstfskellfaa	360
QY	402	LNFTQTEKLRGVFGQOREGVILLRASQEOITRELTRDSDSRRHWHLRGGESSRGYPNLFNKR	461
Db	361	lnftqterlrgvlgqgregvillrasqeqireltrdssesrrwhlrggessrgynlfnkr	420
QY	462	PLYSKTKQOAEVNRPEOYRQJQDMDLVYFLANVTQGSMMGFTRSTKVVVVAASGADV	521
Db	421	plysktkyqaevnrpedytrjqldmdvsvflantltqgsmmgfflntrstkvvaasgadv	480
QY	522	EMACPHLTSGRHSGRGSGGRKRIEEDEDVHYEOVRALSREAIIVLAGHPVYVSSGNNLL	581
Db	481	emacphltsgrlrgsggrkriheeeevhyeyralskreaiivlaaghpvfvssgennll	540
QY	582	LEAFGINAQNHNENFLAGRENNVLOQIETPOMAMELAFAPRKKEVEESFNODOSIFPPGR	641
Db	541	lfafginaqnnhenflagrennvloqilepqsmetelaafaarskveeeltnsgdesilffppgr	600
QY	642	OHQOQSPRSTROQOPVYSILDFVGF	666
Db	601	qhqqgsprstkgqbpvysilldtvgf	625

```

RESULT 4
R20181
ID R20181 standard; Protein; 566 AA.
xx
xx
AC R20181;
xx
xx
DT 16-APR-1992 (first entry)
xx
DE Sequence encoded by 67 kd T. cacao protein cDNA.
xx
xx
KW Cocoa; flavour; vicillin; seed storage protein.
xx
xx
OS Theobroma cacao.
xx
xx
PN WO9119801-A.
xx
xx
PD 26-DEC-1991.
xx
xx
PF 07-JUN-1991; 91WO-GB00914.
xx
xx
PR 11-JUN-1990; 90GB-0013016.
xx
xx
PA (MRSC ) MARS UK LTD.
xx
xx
PI Spencer ME, Hodge R, Deakin EA, Ashton S;
xx
xx
DR WPI: 1992-024418/03.
xx
xx
DR N-PSDB: Q20377.
xx
xx
PT Recombinant cocoa proteins - are responsible for flavour in cocoa
beans and produced in large quantities using yeast and bacterial

```



```

Db      265 hggdkgllivrateeqretlrhaseggshpwb1ppfge--srgpyslldqrpslanqh9 323
OY      471 AYEKVEDYRQLODMPLVFANVTOGSMKGPFFNTRSTKVVVVASGEADVEMACPRLSG 530
Db      324 lyeadaersfhdlehdvsfanitagsmsaplfntrsfklayvpngkyaaevcpbhrs 383
OY      531 RHG----GRGGKRHEEEDVH-----YEOVARLSKREAIIVLAGHPVFEVSSG 576
Db      384 qggeserdkgriseeeseeseegeaaggyhltlrarispglafvypaghpftvavsr 443
OY      577 NEULLFARINNONNHENTLAGERNVLOQIERPAMELFAAPRKVEESFNSQDSIF 636
Db      444 dsnlqivclevhadrnekvlagad-nvllqldrvakalsfaskaeavdevlgsrrekgf 502
OY      637 FPGPRO---HQOOSPRSTKQOQ 655
Db      503 lpgpeesgghreeregeereee 524

RESULT      8
W62837
ID      W62837 standard; Protein: 637 AA.
AC      W62837;
XX
XX
XX      27-OCT-1998 (first entry)
DE      Hordeum vulgare antimicrobial protein.
XX      antimicrobial protein; infestation: control.
XX      Hordeum vulgare.
XX      WO9827805-A1.
XX      02-JUL-1998.
XX      PD
XX      PE      22-DEC-1997; 97WO-AU00874.
XX      PR      20-DEC-1996; 96AU-0004275.
XX
XX      (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
XX      Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX
XX      WPI; 1998-377279/32.
XX      DR
XX      PT      Novel anti-microbial protein from e.g. Macadamia integrifolia -
XX      useful for controlling microbial infestations of plants or mammals
XX
XX      Claim 1; Page 60-62; 96pp; English.
XX      PS
XX      CC      The sequence is that of an antimicrobial protein which can
XX      CC      be used to control microbial infestations in plants and mammalian
XX      CC      animals.
XX
XX      SQ      Sequence 637 AA;

Query Match      24.4%; Score 864.5; DB 19; Length 637;
Best Local Similarity 35.1%; Pred. No. 1.1e-66;
Matches 222; Conservative 98; Mismatches 219; Indels 93; Gaps 20;

OY      70 SKYDNOEDPO--TECOQCORCRQESGPRQOQYCORCKEICEEEREYVNRQDPOQVE 127
Db      27 aashdeddrrgshsqcvcqrger--pr---ysharcvgec-----tddgqgh- 71
OY      128 QCOKHQCRRETERHMQTCQQRCEERYEKEKKQAKRYEQQREDEEKYEERKKEEDNKR 187
Db      72 -----grhegeeggrgrgwhgegereehgrgrgwhgegereehgrgrgwh 119
OY      188 DPQGRREYDCCRRCCEQGRFROHQCOLRCRQEQRQHGRC---GDMANPQRGSGKRYEEG 243

```

```

Db      120 gegeereerggrghrhege-----reeerggrgrghgegereeeegrggrgeg 170
OY      244 E--EESQDN--PYVPERSLSTRFRTEEGHISVLENYGSKILRALKNRLVLEENPN 299
Db      171 erdeegdsrrrvpfprstrllrsghtvralrpdqvsrlllgrldvaimewnprr 230
OY      300 AFVLPHTLADAILLVIGRGALKMTHNDRESYNLECGDVRIRIPAGTFYLIINRDNER 359
Db      231 afvvpgrftddggyvvaqgsvltvlengekrsytkvegvlvapgslmhlanldgrtk 290
OY      360 LHTAKFLQITSTGQKKEFPFAGQNDPEPLSTFSEKILEALNTOTEKRGVFGQO--- 416
Db      291 lvalkllhltsvpqkf-qfl-----svkpllaelskrvlraaflktsderlerlfngrqg 344
OY      417 ---REGVIRASQEOJRELTRDSE---SRHWHIRRGCESSRGPVYNLFNKRPILYSKMGYQ 470
Db      345 ektrsvsivraseeqllreltraeaegqghrwpdlppfrgdsrldllegprkianhgr 404
OY      471 AYEKVEDYRQLODMPLVFANVTOGSMKGPFFNTRSTKVVVVASGEADVEMACPRLSG 530
Db      405 lyeadaersfhlangdvrvavaniitpasmcapylntqsfklavleggevgivcpbl-9 463
OY      531 RHG----GRGGKRHEEED-----VHREOVARLSKREAIIVLA 566
Db      464 reseserehkgrrreeeeddqrqrrrgseseseeeqgyetvrvatrsigatfvpp 523
OY      567 GHPVFEVSS--GNENLLFARINNONNHENTLAGERNVLOQIERPAMELFAAPRKVE 624
Db      524 ghpyveistssggsnlgvvcfelnaernewlagra-nvlygklgsbpageltfgrparev 582
OY      625 EESFNSQDSI-FPGPROHQOOSPRSTKQOQ 655
Db      583 gevfragdqdegfvagp---eqgsregegeqe 611

RESULT      9
Y15244
ID      Y15244 standard; Protein: 626 AA.
XX
XX      Y15244;
AC
XX
XX      09-NOV-1999 (first entry)
DE
XX      Peanut allergen, Ara h 1, amino acid sequence.
XX      DE
XX      allergy; immune response; transgenic; allergen; epitope;
XX      KM      Immunoglobulin E; Ig E; binding site; peanut.
XX      OS
XX      Arachis hypogaea.
XX      PN
XX      WO9938978-A1.
XX      PD
XX      05-AUG-1999.
XX      PE      29-JAN-1999; 99WO-US02031.
XX      PR      27-AUG-1998; 98US-0141220.
XX      PR      31-JAN-1998; 98US-0073283.
XX      PR      13-FEB-1998; 98US-0074590.
XX      PR      13-FEB-1998; 98US-0074624.
XX      PR      13-FEB-1998; 98US-0074633.
XX
XX      PA      (SOSI/) SOSIN H.
XX      PA      (UYAR-) UNIV ARKANSAS.
XX      PA      (UYNY ) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.
XX      PI      Bannon GA, Burks AW, Sampson HA, Sosin H;
XX      WPI; 1999-479189/40.
XX      DR      N-PSDB; 206382.
XX
XX      Modified allergen with reduced IgE binding, useful for treating e.g.
PT

```

Query Match	24.0%	Score 849.5	DB 20	Length 626
Best Local Similarity	34.4%	Pred. No. 2,26-65		
Matches 212	Conservative 108	Mismatches 204	Indels 93	Gaps 20
93	ESGPRQOQY	---CORCKEICEEEYNNRQRPQOQYEQOCQHOREPR	-----H	142
26	kspgpkktkempcaqclscqcpoddlkq	-----acscrctkleydpcrcvdygrh		78
143	MOTQOORC	---ERR-----YEKERKQOAKRYEQOQREDEKYEERMEEDNKRDPQOR		192
79	tgltngtrppgetrtgrpgdyddtr	---qprteegrtwgpagrerereed	---wrgpr	133
193	EVEDCRRCQEQEPKQOHOCOLRCRQOQOHQRGDMMRPQKSGSRYEEGEGSDNY			252
134	---edwrtspbqg-pr	-----kltrpgege	---gewgltprshvretsrnfp	175
253	YFDESLSTRFRTBECHISVLNENFGRSKLALKNYRLVLEANDNATVLTHTLDADAI			312
176	yprstrfstrcrryngnrrivlqrfdqtrsfqnlgnrrivgleakpntclvlpkhdadnl			235
313	LIVIGRGALMKIMHNDRESYMLGCGDVIRIAGTTFVYLINDNMRERLIAKFLQITSP			372
236	lvlgqgaetlvangntrksfnldeghalrpsgfisyllnrhdgnlrvaklmspnvtr			295
373	GOYKEFPFAGGONPEPYLSTFSKEILALNQTQELRGV	-----GQOR		417
296	ggfedtfpssrdqssylqgfsrntleaafeofneirvlleenaagegeorgqrwat			355
418	-----EGYIIASQOERELTRDSESHMHIRGSGESRG	-----PYLFFMKRPLYSNK		467
356	tsenneegylivksxhevelckhaks	-----vskysseeeidlnpnlrgepdlsmn		410
468	YGOAVEVKEPRDR-LQODMDLSVFIANVYQSGMGPFFETRTSTKVVVYVSGADVEMACP			526
411	fgkilevprdkknqplqtdlmmlltvelkegslmrlphrfskmmvlyvynkgfngnlvav			470
527	HLSGRHGGGGKRIHEEDVHYE	---OVR---ARLSKREAIIVLAGHPVVVSSGNEN		579
471	rkeqgrgr	---reeeedeedeegsnevrtyrltkegdvflmpaahpvalnasseh		527
580	ILLFAGFINAONHNFELAGRENNYLOQIEPAMELAFAPRKVEYEDSNQDSITFFFG			639
528	ll--gfglaenhnrlflagdkdnvldqlekakaldlpgsgegyeklikngkeshlvsa			585
640	PROHOOQSPRSTKQOOP			656
586	rpgsgsgpspspeksp			602

```

DE Peanut allergen 1168391 Ara h I protein fragment.
XX
XX Major histocompatibility complex; class II; desensitizing; human;
KM allergen; grass; tree; weed; pollen; fungi; mould; food; insect; stinging
KM chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KM screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KM cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
XX mice; gerbil; vaccine; treatment; prevention; hypersensitivity; peanut.
OS Arachis hypogaea.
XX
XX WO9934826-A1.
PM
XX 15-JUL-1999.
XX
XX 11-JAN-1999; 99WO-GB00080.
XX
XX PF 21-SEP-1998; 98GB-0020474.
XX PR 09-JAN-1998; 98GB-0000445.
XX
XX PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX PI Kay AB, Larche M.
XX WPI, 1999-458255/38.
XX DR
XX PT Desensitizing patients to polypeptide allergens
XX PS Example 6; Page 70-71; 117pp; English.
XX
XX This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC Class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC Class II molecule. The methods can be used for
CC desensitising patients to allergens present in e.g. grass, tree and weed
CC (including rye) pollens, fungi and moulds, foods, stinging insects,
CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
CC Tenebrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC produce immunological vaccines which may be used to prevent and/or treat
CC conditions involving hypersensitivity to allergens. This sequence
CC represents a peanut (Arachis hypogaea) allergen 1168391 Ara h I.
XX
SQ Sequence 626 AA:
Query Match 24.0%; Score 849.5; DB 20; Length 626;
Best Local Similarity 34.4%; Pred. No. 2.2e-65;
Matches 212; Conservative 108; Mismatches 204; Indels 93; Gaps 20.
QY 93 ESGPROOQY---CORCKKEICEEEENORPOOOYEEOCQHQRREPR-----H 142
| | | : : : | : | : | : : | : : | : | : | : | : | : | : |
Db 26 kspgkykkempcaqrclsqcgpeddlkqk-----acesrccklcydpcvcydpgrgh 78
QY 143 MOTCOQR---ERR-----YEKKRKOQKRYEEOQFEDEKYEERMKEEDNKDPQOR 192
| | | | | | : : : | : : | : : | : | : | : | : | : | : |
Db 79 tglqtngirpbgerttgyrpgdyddrr--qprreagrtwpgaprerereed-wrgpr-- 133
QY 193 EVEDCRRCCEOEERQQOQCQLRCQEQQROHORGGDDMMNPQGSGGRRYEEGEDSDMPX 252
| | | | | | | | | : : | : | : | : | : | : | : | : | : |
Db 134 --edwrfrshdq-pl-----kltregtege---gewgtfsgshvireetsrnplf 175
QY 253 YFDEKSLSTRFTTEEGHTSVLENFYGRSKLRLALNNYRLVLLEANPNAFVPTHLDADAI 312
| | | | | : : : | : | : | : | : | : | : | : | : | : | : |
Db 176 yfprffrtfyryngnqgrltrvgtrfdqrsrqfgndqnhtvtqdeakpnltvlpkhadadai 235
QY 313 LLVIIGRGALKMIHHNDRESYNLECGDVIIRIPAGTFFYLINRDNNERLIARFILOTISTP 372
| | | : : : | : | : | : | : | : | : | : | : | : | : | : |
Db 236 lviqggatctvcvanmknksfnldghalrlpsbfisyflnrhdcsnrlrvyaktsmpvto 295

```


